DISEI - Università degli Studi di Firenze

Working Papers - Economics

Endogenous Public Evidence in Committee Persuasion Games and Private Communication

L. Ferrari

Working Paper N. 06/2017

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze, Italia
www.disei.unifi.it

Endogenous Public Evidence in Committee Persuasion Games and Private Communication.

Luca Ferrari*

Abstract

I extend a Jury decision making model allowing jurors to observe, in addition to their private information, public evidence strategically designed by a Prosecutor who wants to maximize the probability of conviction of a defendant under trial. I show that jurors' communication, modeled as a non-binding straw vote before the final verdict, allows the Jury to force Prosecutor to supply evidence whose accuracy becomes perfect as the number of jurors goes to infinity. Thus, the same outcome predicted by the Condorcet's Jury Theorem is reached although evidence is strategically designed so as to prevent this outcome. (JEL D72,D82,D83.)

One of the most common environment in which the Condorcet's Jury Theorem is called to show the superiority of majority voting is that of a criminal trial. Constant in these models is the idea that jurors possess some private information about the payoff relevant state that is crucial to the trial. More precisely, according to Feddersen and Pesendorfer (1998) these pieces of private information reflect jurors' model of reality and stem from the different way in which jurors interpret the evidence publicly shown during the trial. Howbeit, public evidence should move jurors' opinions in the same way were they rational Bayesian decision makers[1]. In this model I introduce strategic public evidence in addition to juror's private information. Under this formulation, jurors may end up having different private opinions about the case under trial although they will interpret public evidence in the same way. Nevertheless, evidence is designed by a Prosecutor in such a way to maximize the probability of conviction – regardless from the innocence of the defendant. This is a simple Bayesian Persuasion

*Department of Economics and Management Via delle Pandette, 9, 50127 Firenze, Italia, email: luca.ferrari.sdc@gmail.com

[1] In addition, in these kind of models, private information a is necessary assumption to study strategic voting behaviors and information aggregation problems.

problem à la Kamenica and Gentzkow (2011) in which a Jury characterized by jurors with endogenous opinion diversity has to take a decision to convict or to acquit the defendant when a biased Prosecutor is unconstrained in his ability to supply strategic persuasive evidence[2]. In general, under fair circumstances, Prosecutor has a lot to gain by strategically designing his evidence. Nevertheless, I also show that, if jurors can communicate their private information as in Coughlan (2000), Prosecutor's strategy then depends on the number of jurors. In particular, as long as jurors' private information is sufficiently informative, communication allow them to reach a very informative common posterior which force Prosecutor to increase the accuracy of his investigation as the number of juror increases. That is to say, in the limit, the Jury reaches a correct verdict almost surely as Predicted by the Condorcet's Jury Theorem.

1 The Model

There are two states of nature $\omega \in \Omega = \{0, 1\}$ where 1 stands for the guilt of the defendant and 0 for his innocence. At the beginning of the game, n jurors privately and independently observe realization $s_i \in \{0, 1\}$ drawn from the common knowledge distribution π: $\pi(s_i = 0|\omega = 0) = \pi(s_i = 1|\omega = 1) = p \in (1/2, 1)$. Although jurors share the same prior belief $\mu(\omega) \in \Delta(\Omega)$ and $\mu(0) > \mu(1)$[3], after the observation of their private realizations they update their beliefs using Bayes rule end they may end up having either posterior belief $\mu_1(\omega) = \frac{\mu_0(\omega)\pi(1|\omega)}{\sum_\omega \mu_0(\omega)\pi(1|\omega)}$ or posterior belief $\mu_0(\omega)$ defined analogously. Importantly, since $p > 1/2$, it follows that $\mu_1(1) > \mu_0(1)$ and $\mu_0(0) > \mu_1(0)$. Then, Prosecutor, who is not informed neither about each juror's private realization nor ω, designs an investigation $\hat{\pi}$ that consists of two conditional probability distributions, $\hat{\pi}(0|0)$ and $\hat{\pi}(1|1)$. His objective is that of maximizing the probability of conviction. Then, some public realization (evidence) $\hat{s} \in \{0, 1\}$ is drawn from $\hat{\pi}$ and jurors update their belief using Bayes rule[4]. I will denote the so reached posterior belief

[2]This model is also closely related to Alonso and Câmara (2016) and Wang (2015) who study a similar problem in which individuals differ in their preferences. Nevertheless, here, individuals' differences are information based, endogenously generated within the model and unknown to sender.

[3]This assumption is made for the sake of exposition, otherwise Prosecutor may have no interest in providing additional evidence.

[4]Implicit in this model is the idea that Prosecutor can commit to truthfully reveal the outcome of the investigation. Although this might constitute a strong assumption, as Kamenica and Gentzkow (2011) underline, Prosecutor is required by Law to truthfully reveal the oucome of the investigation and this prevents cheap-talk equilibria to arise. See Sobel (2010) for a related discussion.

with $\mu_{s,\hat{s}}(\omega)$. Then, jurors cast a vote $x \in \{A, C\}$ to acquit or to convict the defendant and a verdict is made using majority rule. After, payoffs are realized. Prosecutor gets 1 whenever he obtains a conviction, 0 otherwise. Jurors' payoff is 0 whenever their verdict matches the state of Nature, $-q \in (1/2, 1)$ when an innocent defendant is convicted and $-(1-q)$ when a guilty defendant is acquitted.

I consider two cases: in the first, jurors can not communicate their private information, in the second they are allowed to do that reaching the common posterior belief $\mu_{\mathbf{s},\hat{s}}$ where $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ denotes the realized profile of private information.

I focus the attention on Sender-preferred subgame perfect equilibrium (see Kamenica and Gentzkow (2011)) which require that, whenever jurors are indifferent at a given belief, they play sender's preferred action and I restrict the attention to obedient voting strategies in which jurors neglect their private information and vote according to the observed public evidence. In equilibrium, obedience requires that, for all $s_i \in \{0, 1\}$:

$$U(C|s_i, 1) \geq U(A|s_i, 1) \text{ and } U(A|s_i, 0) > U(C|s_i, 0)$$

where $U(\cdot|s_i, \hat{s})$ denotes jurors' expected payoff[5].

1.1 No Communication

To begin with, notice that there are many voting strategies that could be considered. In this note I restrict the attention on the case in which jurors vote according to the observation of the public evidence[6]. Notice that this obedient behavior clearly constitutes an equilibrium of the voting game since no unilateral deviation may change the final verdict, see for instance Austen-Smith and Banks (1996). Prosecutor chooses $\hat{\pi}$ so as to maximize the probability of conviction

$$\Pr(\text{conviction}) \equiv \tau(\hat{\pi}) = \mu_0(1)\hat{\pi}(1|1) + \mu_0(0)(1 - \hat{\pi}(0|0))$$

which is the unconditional probability of showing evidence in favor of conviction $\hat{s} = 1$. Indeed, Prosecutor's expected payoff is exactly $\tau(\hat{\pi})$ since jurors convict the defendant if and only if, they observe $\hat{s} = 1$. There are several ways to

[5]Therefore, in these equilibria the verdict is always unanimous.

[6]Furthermore, voting strategies that make use of jurors' private information may not be optimal from Prosecutor's point of view as this private information tends to aggregate. That is to say, when n is large, most of the jurors will have the correct piece of private information as $p > 1/2$.

solve this problem (see for instance Alonso and Câmara (2016)). However, the simplest is to rely on Kamenica and Gentzkow (2011)'s Propositions 4 and 5 which characterize the optimal investigation:

i) $\hat{\pi}(1|1) = 1$;

ii) Conditional on the realization of $\hat{s} = 1$, jurors must be indifferent between convicting and acquitting the defendant:

$$U(C|1) = U(A|1).$$

Point i) says that the optimal investigation is always truthful when the defendant is actually guilty which is a straightforward point. Indeed, in $\omega = 1$ jurors and Prosecutor's preferences are aligned. Point ii) says that jurors must be indifferent between acquitting or convicting conditional on the observation of public evidence $\hat{s} = 1$. To understand why, notice that now Prosecutor can only move $\hat{\pi}(0|0)$ since $\hat{\pi}(1|1)$ is optimally set to 1. If he sets $\hat{\pi}(0|0)$ equal to 0, then he always report evidence $\hat{s} = 1$, the investigation is completely uninformative and jurors will never vote according to this evidence. However, as long as $\hat{\pi}(0|0)$ increases, evidence $\hat{s} = 1$ becomes increasingly reliable up to a point in which jurors are indifferent between acquitting or convicting the defendant. Prosecutor shall not increase the probability further as it is also clearly increasing the probability of reporting non favorable evidence.

In order to pin down the optimal investigation, recall that at $\hat{s} = 1$, all types of jurors must weakly prefer to convict. First, notice that when $\hat{\pi}(1|1) = 1$, $\mu_{s_i,0}(0) = 1$ from Bayes rule for all $s_i \in \{0,1\}$. Therefore, for all jurors it is always optimal to acquit conditional on $\hat{s} = 0$. On the contrary, since $\mu_0(0) > \mu_1(0)$ it is straightforward to observe that, if at $\hat{\pi}$, type-0 jurors vote obediently, then also type-1 jurors vote obediently[7]. From the definition of payoffs, it follows that $U(C|0,1) = -q\mu_{0,1}(0)$ whereas $U(A|0,1) = -(1-q)\mu_{0,1}(1)$. Since $\hat{\pi}(1|1) = 1$, after some algebra the condition $U(C|0,1) = U(A|0,1)$ becomes

$$\hat{\pi}(0|0) = 1 - \frac{1-q}{q}\frac{1-p}{p}\frac{\mu(1)}{\mu(0)}. \tag{1}$$

Proposition 1. *The optimal investigation persuading a Jury made of privately informed jurors satisfies*

$$\hat{\pi}(1|1) = 1 \textit{ and } \hat{\pi}(0|0) = 1 - \frac{1-q}{q}\frac{1-p}{p}\frac{\mu(1)}{\mu(0)}$$

[7] In particular, type-1 jurors may also accept uninformative investigations which do not change their beliefs if they are already prone to convict.

and the defendant is convicted with probability

$$\tau_1(\hat{\pi}) = \mu(1) + \mu(0)(1 - \hat{\pi}(0|0)) \geq \mu(1).$$

Notice that Prosecutor's expected payoff is larger than the prior $\mu(1)$ which means that persuasion increases his chances of convicting the defendant. On the contrary, if jurors were able to exclusively rely on their private information, then an extremely large jury would convict the defendant with probability close to $\mu(1)$. This underlines Prosecutor's strategic extent to prefer jurors from neglecting their private information. In fact, the optimal investigation does not depend on the number of jurors since jurors can not communicate.

1.2 Communication

In this subsection I allow jurors to truthfully communicate their private information so that, during the communication stage they reach a common posterior belief $\mu_{\mathbf{s},\hat{s}}(\omega)$. As in Coughlan (2000) this may be the result of a non-binding straw-vote in which all jurors vote informatively according to their private signal.

To begin with, observe that there are exactly $n + 1$ profiles of private information that generate different posteriors and that those posteriors can be simply ordered in the number of "yes" to convict[8]. I index those profiles with $\mathbf{j}$ so that $\mu_{\mathbf{j}'}(1) > \mu_{\mathbf{j}}(1)$ for all $\mathbf{j}' > \mathbf{j} \in \{\mathbf{0}, \mathbf{1}, \ldots, \mathbf{n}\}$. Thus, $\mathbf{j}$ denotes the number of jurors who observed $s_i = 1$. Furthermore, π defines two conditional probability distributions over realized profiles of types. Let

$$\rho(\mathbf{j}|\omega) = \binom{\mathbf{n}}{\mathbf{j}} \pi(1|\omega)^{\mathbf{j}} \pi(0|\omega)^{\mathbf{n}-\mathbf{j}}$$

denote the conditional probability of profile $\mathbf{j}$ and $Z(\mathbf{j}|\omega) = \sum_{\mathbf{i}=\mathbf{j}}^{\mathbf{n}} \rho(\mathbf{i}|\omega)$. Now Prosecutor's problem can be approached by finding a cutoff Jury profile $\mathbf{j}$ to target. Intuitively, the optimal investigation rule can be written as

$$\hat{\pi}(0|0) = 1 - \frac{1-q}{q} \frac{\mu_{\mathbf{j}^*}(1)}{\mu_{\mathbf{j}^*}(0)} \tag{2}$$

where $\mathbf{j}^*$ is defined as

$$\mathbf{j}^* = \text{argmax } \mu(0)(1 - \hat{\pi}_j(0|0)) Z(\mathbf{j}|0) + \mu(1) Z(\mathbf{j}|1).$$

Notice that at a solution to this problem exists since the number of profiles is finite. From Prosecutor's point of view, the composition of the Jury, and thus

[8]That is, the number of jurors who observed $s_i = 1$.

the common posterior, is a random variable. Therefore, Prosecutor faces the following tradeoff. If he targets the worst Jury $\mathbf{j} = \mathbf{0}$, the Jury in which all jurors have observed a "no" to convict, he has to supply a very informative investigation although he is sure to get a conviction when $\hat{s} = 1$. On the contrary, he can target a more friendly Jury $\mathbf{j} > \mathbf{0}$ in which some jurors has observed a "yes" to convict. This allows him to reduce the accuracy of the investigation. However, he is no longer sure to get a conviction when $\hat{s} = 1$. The reason is simple. A less friendly Jury may have realized and the information content of the investigation is not reliable enough to induce the Jury to follow Prosecutor's evidence. Nevertheless, when n is large, targeting the worst Jury is not optimal from Prosecutor's point of view. Intuitively, when n jurors observe n "no" to convict, after the communication stage they reach the common posterior

$$\mu_{\mathbf{j}=\mathbf{0}}(1) = \frac{\mu(1)(1-p)^n}{\mu(1)(1-p)^n + \mu(0)p^n}$$

which approaches 0 very fast as $n \to \infty$ thus forcing quickly Prosecutor to raise $\hat{\pi}(0|0)$ close to one. However, the next proposition shows that as $n \to \infty$ Prosecutor can do no better then supplying the fully informative investigation with $\hat{\pi}(0|0) = 1$.

Proposition 2. *When jurors communicate and the number of jurors approaches infinity, Prosecutor's investigation becomes an increasing function of the number of jurors and as $n \to \infty$ the investigation becomes fully informative, that is $\hat{\pi}(0|0) \to 1$.*

Proof. From (1.2) the probability of conviction under communication can be written as

$$\Pr(\text{conviction}) = \mu(0)(1 - \hat{\pi}(0|0))Z(\mathbf{j}|0) + \mu(1)Z(\mathbf{j}|1)$$

which depends exclusively on $\mathbf{j}^*$ since $\hat{\pi}(0|0)$ satisfies (2). However, recall that, conditional on ω, the number of "yes" follows a binomial distribution. Clearly, in state ω the expected proportion of "yes" is $n\pi(1|\omega)/n = \pi(1|\omega)$ whereas the variance of the proportion of "yes" is

$$\frac{np(1-p)}{n^2} = \frac{p(1-p)}{n} \to 0 \text{ as } n \to \infty.$$

It follows that, as $n \to \infty$, in state $\omega = 0$ the Jury will be characterized by the proportion $1 - p$ of "yes" whereas in state $\omega = 1$ the proportion of "yes" tends to p. Therefore, I can neglect the terms $Z(\mathbf{j}|\omega)$ as there will be just two posteriors to target. However, it is clear that Prosecutor must target the Jury characterized by $(1-p)$ "yes", that is $\mathbf{j}^* \to n(1-p)$, and thus common posterior belief

$$\mu_{\mathbf{j}=n(1-p)}(1) = \frac{\mu(1)p^{(1-p)n}(1-p)^{np}}{\mu(1)p^{(1-p)n}(1-p)^{np} + \mu(0)(1-p)^{n(1-p)}p^{np}}$$

which goes clearly to 0 as $n \to \infty$. Intuitively, when n is large, the proportion of "no" provides overwhelming evidence in favor of state $\omega = 0$. Indeed, from (2), the optimal investigation when jurors hold this belief becomes

$$\hat{\pi}(0|0) = 1 - \frac{\mu(1)p^{(1-p)n}(1-p)^{np}}{\mu(0)(1-p)^{n(1-p)}p^{np}} \frac{1-q}{q}$$

and, finally, the fact that $\lim_{n\to\infty} \hat{\pi}(0|0) = 1$ concludes the proof. $\square$

2 Conclusions

In this article I questioned the idea that jurors' private information stems from different ways through which they may interpret public evidence during a trial. Thus, along with jurors' private information, I introduce public evidence strategically designed by a Prosecutor who wants to maximize the probability of conviction. I showed that, due to the presence of public information, there are equilibria in which Prosecutor increases the probability of conviction by strategically designing his evidence. I also showed that, in these equilibria, communication forces Prosecutor to supply more informative investigations when the number of jurors is large. In addition, in the limit, the Jury makes the correct verdict almost surely achieving the same positive outcome predicted by the Condorcet's Jury Theorem through a different mechanism. To conclude, in this article I showed that communication may increase the chances of making the correct final decision even when public evidence is strategically designed in such a way to sway the Jury towards a conviction. Nevertheless, the game indeed exhibit many other equilibria whose study will be subject of future researches.

Acknowledgements

I would really like to thank Santiago Oliveros, Leonardo Boncinelli and Ennio Bilancini for their support during this research. I am also indebted to Dino Gerardi . All mistakes remain mine.

References

References

Ricardo Alonso and Odilon Câmara. Persuading voters. *The American Economic Review*, 106(11):3590–3605, 2016.

David Austen-Smith and Jeffrey S Banks. Information aggregation, rationality, and the condorcet jury theorem. *American political science review*, 90(01): 34–45, 1996.

Peter J Coughlan. In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *American Political science review*, 94(02):375–393, 2000.

Timothy Feddersen and Wolfgang Pesendorfer. Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political science review*, 92(01):23–35, 1998.

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *The American Economic Review*, 101(6):2590–2615, 2011.

Joel Sobel. Giving and receiving advice. In *Econometric Society 10th World Congress*, 2010.

Yun Wang. Bayesian persuasion with multiple receivers. Technical report, Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, 2015.